63 3-2

RESEARCH REPORT 30
20 DECEMBER 1962
I.E.R. 172-35

SYMMETRIC DUAL NONLINEAR PROGRAMS

by

G. B. Dantzig, E. Eisenberg and R. W. Cottle

OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH

UNIVERSITY OF CALIFORNIA - BERKELEY

SYMMETRIC DUAL NONLINEAR PROGRAMS

by

G. B. Dantzig, E. Eisenberg and R. W. Cottle
Operations Research Center
University of California, Berkeley

20 December 1962 Research Report 30

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83) with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

SYMMETRIC NONLINEAR DUAL PROGRAMS

I. Introduction

Most known results on duality in mathematical programming are very closely related to the conditions under which

$$\text{(M)} \qquad \max_{y \in Y} \min_{x \in X} K(x, y) = \min_{x \in X} \max_{y \in Y} K(x, y)$$

where the sets $X \subset R^n$, $Y \subset R^m$ and the function $K : X \times Y \to R$. As an example, the duality of the two linear programming problems:

$$\text{(1)} \qquad \begin{array}{ll} \text{i)} & x \geq 0 , \quad Ax \geq b , \quad \min cx \\ \text{ii)} & y \geq 0 , \quad yA \leq c , \quad \max yb \end{array}$$

is expressed by the statement (M) with $X = R^n_+$, $Y = R^m_+$ and $K(x, y) = cx + yb - yAx$. The statement (M), or the equivalent saddle-point statement, avoids, among other things, one of the apparent features of many duality formulations, i.e., asymmetry. If, for instance, the primal problem is taken to be

$$\text{(2)} \qquad x \in R^n , \; g_i(x) \leq 0 \quad (i = 1, \ldots, m) , \; \min \; g_0(x) \quad .$$

where $g_0, g_1, \ldots, g_m$ are convex-differentiable functions of $x \in R^n$, then a formal dual is

$$\text{(3)} \qquad x \in R^n , \; u_i \geq 0 \; (i = 1, \ldots, m), \; g_0'(x) - \sum_{i=1}^{m} u_i g_i'(x) = 0 ,$$

$$\max \left[g_0(x) - \sum_{i=1}^{m} u_i g_i(x) \right] ,$$

where $g_i'(x)$ is the gradient of g_i at x. Apparently, the dual problem (3) is not, in general, of the same form as the primal problem (2). Furthermore, while the objective function in (2) is assumed convex, the objective function in (3) may be neither convex nor concave, indicating again a lack of symmetry. In [3] one of the co-authors of this note discussed the situation when in (M) $X = R_+^n$, $Y = R_+^m$, and K is given by:

$$K(x, y) = f(x) + g(y) - yAx \tag{4}$$

with f, g convex-homogeneous and f, g, A further restricted by a "feasibility" condition. Results of Fenchel [5] (which may also be found in [7, p. 227]) pertain to (M) with K as in (4) with f, g convex and X, Y closed convex sets. The main result of this paper is what we believe to be a new duality formulation. It can also be very profitably used in investigating general conditions on X, Y and K under which (M) holds.

Additional or related efforts in the area of duality are noted in the bibliography; they include Wolfe's duality theorem [11, Thm. 2] which states that if the constraints in (2) satisfy the Kuhn-Tucker qualification [8, p. 483] and if x_0 solves (2) then there is a $u = (u_1, \ldots, u_m)$ such that the pair (x_0, u) solve (3) With the exception of special cases, such as quadratic programming, the converse of Wolfe's theorem, as treated independently by Huard [6] and Mangasarian [9], requires stronger assumptions. In this paper we present a pair of symmetric dual programming problems related to the statement (M). In a way, this represents an extension of results obtained by one of this paper's co-authors in [1]. In section 2 we connect the results in [1] with the duality theorem which is established in section 3.

The work of Dorn [2] on quadratic programming exemplifies the fact that the formulation of section 1 for primal and dual problems is asymmetric. Dorn considers the primal problem (the sign of his b has been changed):

$$Ax + b \geq 0 , \quad x \geq 0 , \qquad \min(\tfrac{1}{2} x^T Cx + p^T x) , \tag{5}$$

and its dual (with a slight change in notation)

$$-A^T y + Cx + p \geq 0 , \quad y \geq 0 , \quad \max(-\tfrac{1}{2} x^T Cx - b^T y) , \tag{6}$$

where C is a symmetric positive semi-definite matrix. It is shown in [1] that there is a pair of naturally symmetric problems related by duality. They are:

PRIMAL PROBLEM

$$\begin{array}{lll} & \text{Minimize} & \tfrac{1}{2} y^T Dy + \tfrac{1}{2} x^T Cx + p^T x \\ (7) & \text{subject to} & Dy + \quad Ax + b \geq 0 \\ & & x \in R_+^n , \quad y \in R_+^m \end{array}$$

and

DUAL PROBLEM

$$\begin{array}{lll} & \text{Maximize} & -\tfrac{1}{2} y^T Dy - \tfrac{1}{2} x^T Cx - b^T y \\ (8) & \text{subject to} & -A^T y + \quad Cx + p \geq 0 \\ & & x \in R_+^n , \quad y \in R_+^m \end{array}$$

where both C and D are symmetric positive semi-definite matrices. In particular, if C and D are identically zero, then (7) and (8) reduce simply to symmetric dual linear programs.

II. The General Symmetric Duality Statement

Let $K(x, y)$ be a real-valued and twice continuously differentiable function defined on an open subset of $R^{n+m} = R^n \times R^m$ of the form $U \times V$ where U is open in R^n and V is open in R^m. We define $K_1(x, y)$ to be the n-component column vector which is the gradient of K with respect to the x-variable at the point (x, y); similarly, $K_2(x, y)$ is the m-component column vector representing the gradient of K with respect to y at the point $(x, y) \in U \times V$. We introduce a double subscript notation for the matrices of second partials. Thus,

$$K_1(x, y) = \left[\frac{\partial K(x,y)}{\partial x_1}, \frac{\partial K(x,y)}{\partial x_2}, \ldots \frac{\partial K(x,y)}{\partial x_n}\right]^T$$

$$K_2(x, y) = \left[\frac{\partial K(x,y)}{\partial y_1}, \frac{\partial K(x,y)}{\partial y_2}, \ldots \frac{\partial K(x,y)}{\partial y_m}\right]^T$$

$$K_{11}(x, y) = \left(\frac{\partial^2 K(x,y)}{\partial x_i \partial x_j}\right) \; ; \quad K_{12}(x, y) = \left(\frac{\partial^2 K(x,y)}{\partial x_i \partial y_j}\right) \; ;$$

$$K_{21}(x, y) = \left(\frac{\partial^2 K(x,y)}{\partial y_i \partial x_j}\right) \; ; \quad K_{22}(x, y) = \left(\frac{\partial^2 K(x,y)}{\partial y_i \partial y_j}\right) .$$

The dual programming problems we are primarily concerned with in this paper may be stated in general as:

PRIMAL (P)	DUAL (P^*)
$\underset{x,y}{\text{Min}} \,[K(x,y) - y^T K_2(x,y)]$ Subject to: $K_2(x,y) \leq 0$ $x \in R^n_+ \,,\; y \in R^m_+$	$\underset{x,y}{\text{Max}} \,[K(x,y) - x^T K_1(x,y)]$ Subject to: $K_1(x,y) \geq 0$ $x \in R^n_+ \,,\; y \in R^m_+$

as a special case, with K defined by

$$K(x,y) = \frac{1}{2}x^TCx + p^Tx - \frac{1}{2}y^TCy - b^Ty - y^TAx \tag{10}$$

the primal and dual problems above reduce to the quadratic case (7) and (8).

The general strict case will consist of the two problems (P) and (P^*) with K having the following properties:

(i) K is real valued on the cartesian product $U \times V$ (briefly, $K: U \times V \to R$) where U and V are open subsets of R^n and R^m respectively, such that $R_+^n \subset U$, $R_+^m \subset V$.

(ii) K is twice continuously differentiable on $U \times V$.

(iii) for each fixed $x \in R_+^n$, K is strictly concave in y.

(iv) for each fixed $y \in R_+^m$, K is strictly convex in x.

Observe that if we omit "strictly" in conditions (iii) and (iv) then with K as given by (10) for the quadratic case, all the other conditions are satisfied. The general strict case and the quadratic case are, of course, not mutually exclusive; specifically, the quadratic case becomes an instance of the other providing C and D are both positive definite (hence, nonsingular).

We find it convenient to use a single symbol for the objective functions in (P) and (P^*):

$$\begin{aligned} \xi(x,y) &= K(x,y) - y^TK_2(x,y) \\ \eta(x,y) &= K(x,y) - x^TK_1(x,y) \end{aligned} \tag{11}$$

Also, we denote by P the set of all pairs (x,y) satisfying the constraints of (P), i.e., $x \in R_+^n$, $y \in R_+^m$, $K_2(x,y) \leq 0$; similarly, P^* denotes the set of all pairs (x,y) satisfying $x \in R_+^n$, $y \in R_+^m$, and $K_1(x,y) \geq 0$.

In the next section we state precisely what the duality relation between (P) and (P^*) is. At this point, it should be remarked that neither of the problems (P) and (P^*) need be a convex programming problem. For example, in (P) the objective or the constraint functions may fail to be convex, and analogously in (P^*). Furthermore, it should be noted that the duality requirements defined in the next section are of the type which will not hold unless some regularity conditions are imposed on the function $K(x, y)$. In view of the possible absence of convexity (or concavity), the Kuhn-Tucker theorem [8, p. 84] on necessary conditions of optimality seems appropriate to obtain the desired duality relations. Thus, it appears that the constraint qualification (see Appendix) is a suitable regularity assumption. Indeed, this assumption is sufficient.

III. Duality

We shall say that a relation of duality holds between (P) and (P^*) providing that the following two conditions hold:

(a) $$\sup_{(x,y)\in P^*} \eta(x,y) \leq \inf_{(x,y)\in P} \xi(x,y)$$

and

(b) (P) is solvable if, and only if, (P^*) is solvable, in which case

$$\max_{P^*} \eta = \min_{P} \xi$$

(Solvable means "has an optimal solution.")

REMARK: the further condition:

(c) If exactly one of the programs is feasible then its objective function is unbounded in the direction of optimization;

(Feasible means that there exist variables satisfying its constraints)

which holds for dual quadratic or linear programs (i.e., in the quadratic case, see [1, theorem 3] need not hold in the strict case.[/] This may be seen from the example: $K(x,y) = e^x - e^{-y}$; in this case (P^*) is trivially feasible while (P) is infeasible; however $\eta(x,y) = e^x - e^{-y} - xe^x$ which is bounded above for $(x,y) \geq 0$. We shall show that in the strict case[/] a relation of duality holds between (P) and (P^*) providing they satisfy a constraint qualification; we begin by showing (a):

THEOREM 1: $\sup_{P^*} \eta \leq \inf_{P} \xi$

[/] i.e., in the general strict case.

Proof: Using the convention

$$\sup_{P^*} \eta = -\infty \text{ if } P^* \text{ is empty}$$

$$\inf_{P} \xi = +\infty \text{ if } P \text{ is empty}$$

we need only show that if $(x, y) \in P$ and $(\bar{x}, \bar{y}) \in P^*$ then $\eta(\bar{x}, \bar{y}) \leq \xi(x, y)$. By definition of P and P^* we have

$$\text{(12)} \qquad \begin{aligned} &K_2(x, y) \leq 0, \quad x \geq 0, \quad y \geq 0 \\ &K_1(\bar{x}, \bar{y}) \geq 0, \quad \bar{x} \geq 0, \quad \bar{y} \geq 0 . \end{aligned}$$

Thus:

$$\text{(13)} \qquad \bar{y}^T K_2(x, y) \leq 0, \quad x^T K_1(\bar{x}, \bar{y}) \geq 0$$

and

$$\text{(14)} \qquad \bar{y}^T K_2(x, y) - x^T K_1(\bar{x}, \bar{y}) \leq 0 .$$

Since K is convex-concave and differentiable we have (see [8] Vol. 1, p. 405, No. vii)

$$\text{(15)} \qquad \begin{aligned} &x^T K_1(\bar{x}, \bar{y}) - \bar{x}^T K_1(\bar{x}, \bar{y}) \leq K(x, \bar{y}) - K(\bar{x}, \bar{y}) \\ &y^T K_2(x, y) - \bar{y}^T K_2(x, y) \leq K(x, y) - K(x, \bar{y}) \end{aligned}$$

Adding the two inequalities of (15), rearranging terms, and using (14), we obtain:

$$\begin{aligned} \eta(\bar{x}, \bar{y}) = K(\bar{x}, \bar{y}) - \bar{x}^T K_1(\bar{x}, \bar{y}) &\leq \\ &\leq K(x, y) - y^T K_2(x, y) + \left[\bar{y}^T K_2(x, y) - x^T K_1(\bar{x}, \bar{y}) \right] \leq \\ &\leq K(x, y) - y^T K_2(x, y) = \xi(x, y) , \end{aligned}$$

completing the proof of Theorem 1.

NOTE: The proof of Theorem 1 does not require the "strictness" assumption of (iii), (iv); only convexity-concavity and differentiability are essential.

From Theorem 1 follows the obvious remark that if $(x_0, y_0) \in P$, $(\bar{x}_0, \bar{y}_0) \in P^*$ have the property that $\xi(x_0, y_0) \leq \eta(\bar{x}_0, \bar{y}_0)$ then (x_0, y_0) and $(\bar{x}_0, \bar{y}_0)$ are optimal solutions of P and P^* respectively. This will be used in proving:

THEOREM 2 (Duality Theorem): Assume that P and P^* satisfy the Kuhn-Tucker constraint qualification. Then (P) is solvable if, and only if, (P^*) is solvable, in which case

$$\max_{P^*} \eta = \min_{P} \xi$$

Proof: We shall show that if $(\bar{x}, \bar{y})$ solves (P^*) then $(\bar{x}, \bar{y})$ also solves (P) and $\xi(\bar{x}, \bar{y}) = \eta(\bar{x}, \bar{y})$ (the proof of the converse is analogous). Consider the function:

$$\psi(x, y, u) = K(x, y) - x^T K_1(x, y) + u^T K_1(x, y)$$

By the Kuhn-Tucker optimality conditions [8, theorem 1], and since $(\bar{x}, \bar{y})$ solves (P^*), there exists a vector $\bar{u} \in R^n$ satisfying the following:

(16) $$\psi_1(\bar{x}, \bar{y}, \bar{u}) = K_{11}(\bar{x}, \bar{y})(\bar{u} - \bar{x}) \leq 0$$

(17) $$\psi_2(\bar{x}, \bar{y}, \bar{u}) = K_2(\bar{x}, \bar{y}) + K_{21}(\bar{x}, \bar{y})(\bar{u} - \bar{x}) \leq 0$$

(18) $$\psi_3(\bar{x}, \bar{y}, \bar{u}) = K_1(\bar{x}, \bar{y}) \geq 0$$

(19) $$\bar{x}^T \psi_1(\bar{x}, \bar{y}, \bar{u}) + \bar{y}^T \psi_2(\bar{x}, \bar{y}, \bar{u}) =$$
$$= \bar{x}^T K_{11}(\bar{x}, \bar{y})(\bar{u} - \bar{x}) + \bar{y}^T K_2(\bar{x}, \bar{y}) + \bar{y}^T K_{21}(\bar{x}, \bar{y})(\bar{u} - \bar{x}) =$$
$$= 0$$

$$(20) \qquad \bar{u}^T K_1(\bar{x},\bar{y}) = 0$$

$$(21) \qquad \bar{u} \geq 0, \quad \bar{x} \geq 0, \quad \bar{y} \geq 0 \ .$$

It then follows that:

$$(22) \qquad \bar{x}^T K_{11}(\bar{x},\bar{y})(\bar{u} - \bar{x}) = 0$$

$$(23) \qquad \bar{y}^T K_2(\bar{x},\bar{y}) + \bar{y}^T K_{21}(\bar{x},\bar{y})(\bar{u} - \bar{x}) = 0 \ .$$

From (16) and (21) we get:

$$(24) \qquad \bar{u}^T K_{11}(\bar{x},\bar{y})(\bar{u} - \bar{x}) \leq 0$$

and thus:

$$(25) \qquad (\bar{u} - \bar{x})^T K_{11}(\bar{x},\bar{y})(\bar{u} - \bar{x}) \leq 0 \ .$$

But because K is strictly convex in x, it follows that K_{11} is a positive definite matrix, thus $\bar{u} - \bar{x} = 0$, i.e.,

$$(26) \qquad \bar{u} = \bar{x}$$

It then follows from (17) and (26) that the pair $(\bar{x},\bar{y})$ is feasible for (P). From (26) and (19) it follows that

$$(27) \qquad \bar{y}^T K_2(\bar{x},\bar{y}) = 0 \ ,$$

while from (26) and (20) we obtain:

$$(28) \qquad \bar{x}^T K_1(\bar{x},\bar{y}) = 0 \ .$$

We conclude that $(\bar{x},\bar{y}) \in P \cap P^*$ and $\xi(\bar{x},\bar{y}) = \eta(\bar{x},\bar{y})$. Q. E. D.

We state two immediate corollaries:

COROLLARY 1:

If either program (P) or (P^*) is solvable then there is a joint solution (i.e., a pair solving both programs).

Proof: It is clear from the proof of Theorem 2 that in the general strict case the optimal solutions of (P) and (P^*) are the same. For the quadratic case see [1, Theorem 4].

COROLLARY 2:

If $(\bar{x}, \bar{y})$ is a joint solution of (P) and (P^*) then $\bar{y}^T K_2(x, y) = 0 = x^T K_1(\bar{x}, \bar{y})$.

Proof: For the general strict case note Equations (27) and (28); for the quadratic case see the remark after Theorem 4 in [1].

APPENDIX

The Kuhn-Tucker Constraint Qualification

Suppose $F(x) = [f_1(x), f_2(x), \ldots, f_m(x)]^T$ where each f_i is an everywhere differentiable function of $x \geq 0$. Let x^o be such a point. Taking the partial derivatives of the functions f_i at x^o, we define

$$F^o = \left[\frac{\partial f_i}{\partial x_j}\right]^o ,$$

which is regarded as an $m \times n$ matrix, the i^{th} row of which is the gradient (transposed) of the function f_i evaluated at x^o.

Let x^o be a boundary point of the (constraint) set $C = \{x \epsilon R_+^n \mid F(x) \geq 0\}$. Separate the inequalities $F(x^o) \geq 0$, $Ix^o \geq 0$ (where I is the $m \times n$ identity matrix) into

$$F_1(x^o) = 0 \quad , \quad I_1 x^o = 0 \quad , \quad F_2(x^o) > 0 \quad , \quad I_2 x^o > 0 \quad .$$

That is, F_1 consists of the subset of functions f_i in F which vanish at x^o, and F_2 consists of those which are positive at x^o. I_1 and I_2 are diagonal matrices having ones and zeros as diagonal entries. In particular, I_1 has ones in those rows where the components of x^o are zero, and has zeros on the rest of the diagonal. I_2 has ones in the rows where x^o has positive components and has zeros on the remainder of its diagonal.

The Maximum Problem of Kuhn and Tucker [8, p. 483] is: Maximize a differentiable function $g(x)$ constrained by $F(x) \geq 0$, $x \geq 0$. Further restrictions on the constraint set are required in order to make certain conclusions about an optimal solution of the maximum problem. These restrictions are embodied in the constraint qualification which we now state.

For every boundary point x^o of C, any vector differential dx satisfying the homogeneous linear inequalities

$$F_1^o dx \geq 0 \quad , \quad I_1 dx \geq 0$$

is tangent to a differentiable arc contained in C. That is, to every dx satisfying the above inequalities, there corresponds an arc $x = a(\theta)$, $0 \leq \theta \leq 1$, with $x^o = a(0)$, such that $[da/d\theta] = \lambda dx$ for some positive scalar λ.

REFERENCES

[1] Cottle, Richard W., "Symmetric Dual Quadratic Programs," Operations Research Center, University of California, Berkeley RR 19, May 1962.

[2] Dorn, W.S., "Duality in Quadratic Programming," Quart. Appl. Math., Vol. 18, 1960. pp. 155-162.

[3] Eisenberg, E., "Duality in Homogeneous Programming," Proc. Amer. Math. Soc., Vol. 12, 1961. pp. 783-787.

[4] Fan, Ky, I. Glicksberg and A. J. Hoffman, "Systems of Inequalities Involving Convex Functions," Proc. Amer. Math. Soc. Vol. 8, 1957. pp. 617-622.

[5] Fenchel, W., "Convex Cones, Sets and Functions," Princeton University Lecture Notes, Spring 1953.

[6] Huard, P., "Dual Programs," IBM Journal of Research and Development, Vol. 6, 1962. pp. 137-139.

[7] Karlin, Samuel, Mathematical Methods and Theory in Games, Programming and Economics, Addison-Wesley, 1959. Vols. I-II.

[8] Kuhn, H. W. and A. W. Tucker, "Nonlinear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1951, pp. 481-492.

[9] Mangasarian, O. L., "Duality in Nonlinear Programming," Quart. Appl. Math., Vol. 20, 1962. pp. 300-302.

[10] Wolfe, P., "A Duality Theorem for Nonlinear Programming," Quart. Appl. Math., Vol. 19, 1961. pp. 239-244.

BASIC DISTRIBUTION LIST
FOR UNCLASSIFIED TECHNICAL REPORTS

Head, Logistics and Mathematical
Statistics Branch
Office of Naval Research
Washington 25, D.C.

C.O., ONR Branch Office
Navy No. 100, Box 39, F.P.O.
New York City, New York

ASTIA Document Service Center
Arlington Hall Station
Arlington 12, Virginia

Institute for Defense Analyses
Communications Research Div.
von Neumann Hall
Princeton, New Jersey

Technical Information Officer
Naval Research Laboratory
Washington 25, D.C.

C.O., ONR Branch Office
346 Broadway, New York 13, NY
Attn: J. Laderman

C.O., ONR Branch Office
1030 East Green Street
Pasadena 1, California
Attn: Dr. A.R. Laufer

Bureau of Supplies and Accounts
Code OW, Dept. of the Navy
Washington 25, D.C.

Professor Russell Ackoff
Operations Research Group
Case Institute of Technology
Cleveland 6, Ohio

Professor Kenneth J. Arrow
Serra House, Stanford University
Stanford, California

Professor G.L. Bach
Carnegie Institute of Technology
Planning and Control of Industrial
Operations, Schenley Park
Pittsburgh 13, Pennsylvania

Professor A. Charnes
The Technological Institute
Northwestern University
Evanston, Illinois

Professor L. W. Cohen
Math. Dept., University of Maryland
College Park, Maryland

Professor Donald Eckman
Director, Systems Research Center
Case Institute of Technology
Cleveland, Ohio

Professor Lawrence E. Fouraker
Department of Economics
The Pennsylvania State University
State College, Pennsylvania

Professor David Gale
Dept. of Math., Brown University
Providence 12, Rhode Island

Dr. Murray Geisler
The RAND Corporation
1700 Main Street
Santa Monica, California

Professor L. Hurwicz
School of Business Administration
University of Minnesota
Minneapolis 14, Minnesota

Professor James R. Jackson
Management Sciences Research
Project, Univ. of California
Los Angeles 24, California

Professor Samuel Karlin
Math. Dept., Stanford University
Stanford, California

Professor C.E. Lemke
Dept. of Mathematics
Rensselaer Polytechnic Institute
Troy, New York

Professor W.H. Marlow
Logistics Research Project
The George Washington University
707 - 22nd Street, N.W.
Washington 7, D.C.

Professor Oskar Morgenstern
Economics Research Project
Princeton University
92 A Nassau Street
Princeton, New Jersey

BASIC DISTRIBUTION LIST FOR UNCLASSIFIED TECHNICAL REPORTS

Professor R. Radner
Department of Economics
University of California
Berkeley, California

Professor Stanley Reiter
Department of Economics
Purdue University
Lafayette, Indiana

Professor Murray Rosenblatt
Department of Mathematics
Brown University
Providence 12, Rhode Island

Mr. J. R. Simpson
Bureau of Supplies and Accounts
Navy Department (Code W31)
Washington 25, D.C.

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Lincoln Hall, Cornell University
Ithaca 1, New York